

**IMPLEMENTASI ALGORITMA LINEAR REGRESSION UNTUK
PREDIKSI EMISI KARBON DIOKSIDA PADA KENDARAAN MOBIL**



SKRIPSI

**Muhammad Rifqi Al Hakim
0000050617**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2024**

**IMPLEMENTASI ALGORITMA LINEAR REGRESSION UNTUK
PREDIKSI EMISI KARBON DIOKSIDA PADA KENDARAAN MOBIL**



Muhammad Rifqi Al Hakim

00000050617

UMMN

UNIVERSITAS

MULTIMEDIA

NUSANTARA

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2024

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Muhammad Rifqi Al Hakim
Nomor Induk Mahasiswa : 00000050617
Program Studi : Informatika

Skripsi dengan judul:

Implementasi Algoritma Linear Regression Untuk Prediksi Emisi Karbon Dioksida Pada Kendaraan Mobil

merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/ penyimpangan, baik dalam pelaksanaan Skripsi maupun dalam penulisan laporan Skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas akhir yang telah saya tempuh.

Tangerang, 22 Mei 2024



(Muhammad Rifqi Al Hakim)

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

HALAMAN PENGESAHAN

Skripsi dengan judul

IMPLEMENTASI ALGORITMA LINEAR REGRESSION UNTUK PREDIKSI EMISI KARBON DIOKSIDA PADA KENDARAAN MOBIL

oleh

Nama : Muhammad Rifqi Al Hakim
NIM : 00000050617
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Kamis, 30 Mei 2024

Pukul 13.00 s/d 15.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

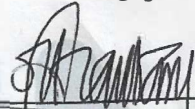
Ketua Sidang



(Eunike Endariahna Surbakti, S.Kom.,
M.T.I.)

NIDN: 0322099401

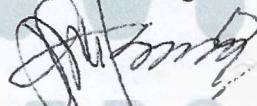
Penguji



(Aditiyawan, S.Komp., M.Si.)

NIDN: 8994550022

Pembimbing



(Adhi Kushadi, S.T., M.Si)

NIDN: 303037304

PJS Ketua Program Studi Informatika,



(Dr. Eng. Niki Prastomo, S.T., M.Sc.)

NIDN: 0419128203

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Muhammad Rifqi Al Hakim
NIM : 00000050617
Program Studi : Informatika
Jenjang : S1
Jenis Karya : Skripsi

Menyatakan dengan sesungguhnya bahwa:

- Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya di repositori Knowledge Center, sehingga dapat diakses oleh Civitas Akademika/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial dan saya juga tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.
- Saya tidak bersedia karena dalam proses pengajuan untuk diterbitkan ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*)**.

Tangerang, 22 Mei 2024

Yang menyatakan



Muhammad Rifqi Al Hakim

U M M N
U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

** Jika tidak bisa membuktikan LoA jurnal/HKI selama enam bulan ke depan, saya bersedia mengizinkan penuh karya ilmiah saya untuk diunggah ke KC UMN dan menjadi hak institusi UMN.

Halaman Persembahan / Motto

”Pay no mind to those who talk behind your back it simply means that you are two steps ahead”

Tupac Shakur

UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: Implementasi Algoritma Linear Regression Untuk Prediksi Emisi Karbon Dioksida Pada Kendaraan Mobil dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara dan juga Pjs. Ketua Program Studi Informatika Universitas Multimedia Nusantara.
3. Bapak Adhi Kusnadi, S.T., M.Si, sebagai Pembimbing pertama yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya tesis ini.
4. Orang Tua, teman dan keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tesis ini.

Semoga skripsi ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 22 Mei 2024



Muhammad Rifqi Al Hakim

IMPLEMENTASI ALGORITMA LINEAR REGRESSION UNTUK PREDIKSI EMISI KARBON DIOKSIDA PADA KENDARAAN MOBIL

Muhammad Rifqi Al Hakim

ABSTRAK

Emisi karbon dioksida (CO_2) mengalami peningkatan terus menerus hingga meningkatkan pemanasan global dan salah satu penyebabnya adalah produksi emisi karbon dioksida (CO_2) dari kendaraan roda empat. Algoritma *Linear Regression* digunakan untuk memprediksi emisi karbon dioksida (CO_2) yang dihasilkan dalam kendaraan mobil dengan variabel yang meliputi ukuran mesin, jumlah silinder, konsumsi bahan bakar, dan emisi karbon dioksida (CO_2) yang dihasilkan. Hasilnya menunjukkan bahwa skor RMSE untuk *training* dan *testing* adalah 13.98 dan 15.12 yang menandakan bahwa pemodelan algoritma *Linear Regression* tidak mengalami *overfitting*, skor R^2 pada hasil *training* dan setelah dilakukan *cross-validation 10-folds* sebesar 0.91 dan 0.909 yang menandakan bahwa pemodelan algoritma *Linear Regression* untuk melakukan prediksi emisi karbon dioksida (CO_2) mampu memprediksi dengan baik. Kesimpulan dari penelitian ini adalah setelah melewati enam tahapan, pemodelan algoritma *Linear Regression* untuk prediksi emisi karbon dioksida (CO_2) berhasil dilakukan dengan akurasi yang mencapai 0.91 dan 0.909 setelah melakukan *cross-validation 10-folds*.

Kata kunci: *cross-validation*, emisi karbon dioksida, kendaraan mobil, *linear regression*, prediksi

UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

***Linear Regression Algorithm Implementation for Carbon Dioxide Emissions
Prediction in Car***

Muhammad Rifqi Al Hakim

ABSTRACT

Carbon dioxide (CO₂) emissions have increased continuously to increase global warming and one of the causes is the production of carbon dioxide (CO₂) emissions from four-wheeled vehicles. Linear Regression algorithm is used to predict carbon dioxide (CO₂) emissions produced in automobile vehicles with variables including engine size, number of cylinders, fuel consumption, and carbon dioxide (CO₂) emissions produced. The results show that the RMSE scores for training and testing are 13.98 and 15.12 which indicates that the Linear Regression algorithm modeling does not experience overfitting, the R² score on the training results and after 10-folds cross-validation is 0.91 and 0.909 which indicates that the Linear Regression algorithm modeling to predict carbon dioxide (CO₂) emissions is able to predict well. The conclusion of this research is that after passing through six stages, modeling the Linear Regression algorithm for predicting carbon dioxide (CO₂) emissions was successfully carried out with an accuracy that reached 0.91 and 0.909 after cross-validation 10-folds.

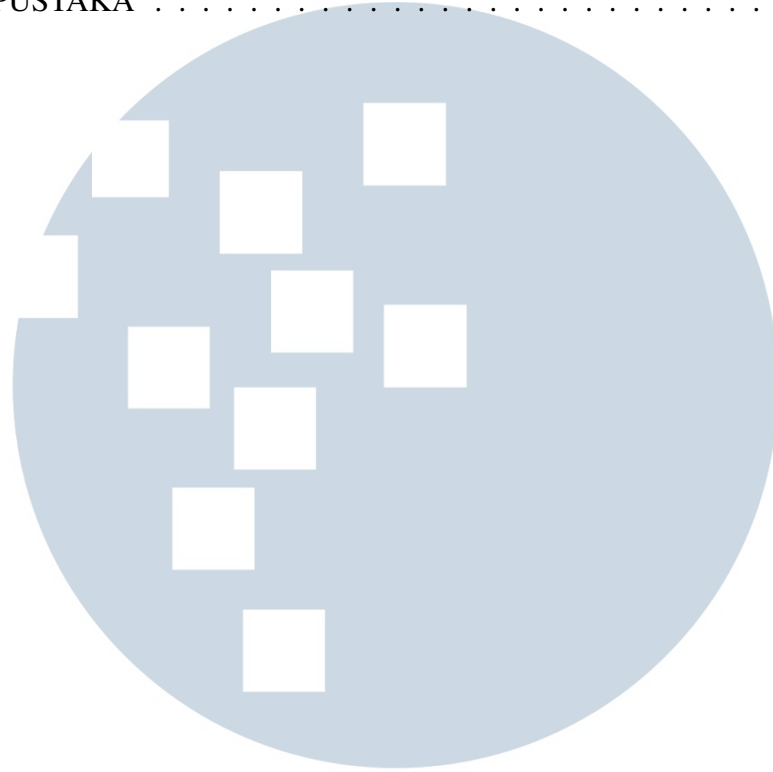
Keywords: car vehicle, carbon dioxide emissions, cross-validation, linear regression, prediction



DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
DAFTAR KODE	xiii
DAFTAR LAMPIRAN	xiv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Permasalahan	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	6
2.1 Emisi Karbon Dioksida (CO_2)	6
2.2 Kendaraan Mobil dan Faktor	7
2.3 <i>Machine Learning</i>	7
2.4 Algoritma <i>Linear Regression</i>	7
2.5 <i>Preprocessing</i>	8
2.6 <i>Root Mean Square Error (RMSE)</i>	9
2.7 Koefisien Determinasi (R^2)	9
2.8 Validasi Silang <i>10-Fold</i>	10
BAB 3 METODOLOGI PENELITIAN	11
3.1 Studi Literatur	12
3.2 Pencarian Dataset	12
3.3 <i>Preprocessing</i>	12
3.4 Pengambilan Sampel	13
3.5 Normalisasi Data	13
3.6 <i>Splitting Data Training dan Data Testing</i>	13
3.7 Pengembangan Model Algoritma <i>Linear Regression</i>	14
3.8 Evaluasi dan Dokumentasi	14
BAB 4 HASIL DAN DISKUSI	15
4.1 Eksplorasi Data	15
4.2 Visualisasi Data	18
4.3 <i>Data Cleaning</i> dan Korelasi Variabel	22
4.4 Pengambilan Sampel	26
4.5 Normalisasi Data	28
4.6 <i>Splitting Data Training dan Data Testing</i>	30
4.7 Pemodelan	31
4.8 Evaluasi	32
BAB 5 SIMPULAN DAN SARAN	34

5.1	Simpulan	34
5.2	Saran	34
DAFTAR PUSTAKA		35



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

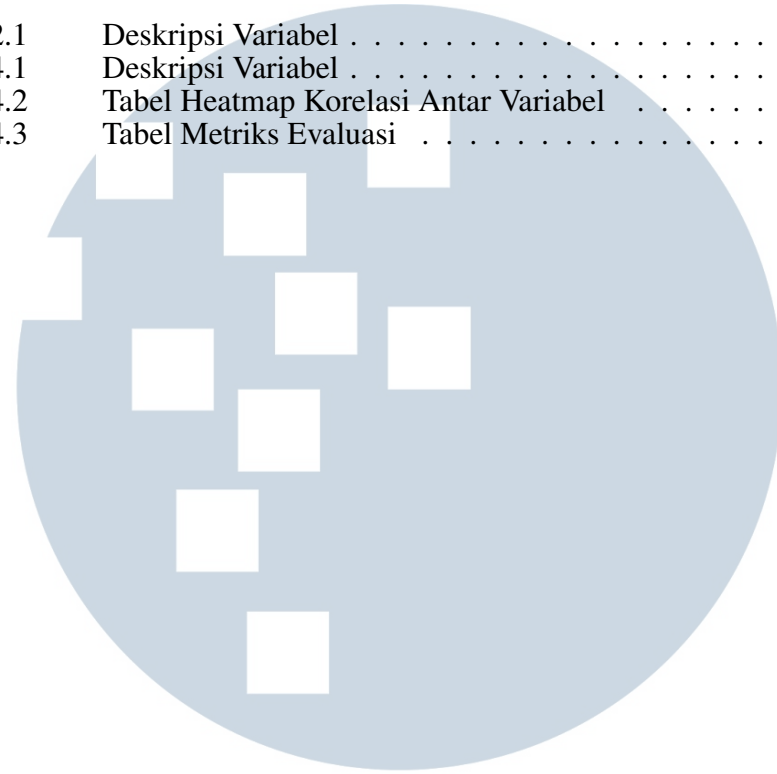
DAFTAR GAMBAR

Gambar 3.1	Flowchart proses pembuatan model	11
Gambar 3.2	Tahap <i>Preprocessing</i>	13
Gambar 3.3	Flowchart proses pembuatan model	14
Gambar 4.1	<i>Import Library</i> di Python	16
Gambar 4.2	Membaca Data Menjadi Tabel	16
Gambar 4.3	Mengecek Nilai Kosong (<i>Null</i>) pada Dataset	17
Gambar 4.4	Mengecek Tipe Data Pada Dataset	18
Gambar 4.5	Visualisasi Variabel Pada Dataset	19
Gambar 4.6	Visualisasi Jumlah Rata-Rata Emisi Karbon Dioksida (CO_2) Pada Kendaraan Mobil Berdasarkan Variabel	20
Gambar 4.7	Identifikasi Data Kendaraan Mobil yang Menggunakan Bahan Bakar Gas Alami	22
Gambar 4.8	Penghapusan Data Kendaraan Mobil yang Menggunakan Bahan Bakar Gas Alami	23
Gambar 4.9	Pembuatan Tabel Untuk Mengecek Korelasi Variabel	23
Gambar 4.10	<i>Plotting</i> Untuk Mengetahui Data <i>Outliers</i>	25
Gambar 4.11	Pengambilan Sampel	27
Gambar 4.12	Pemisahan Sampel Dari Dataset	28
Gambar 4.13	Normalisasi Data Sampel <i>sample.df.Xtest</i>	29
Gambar 4.14	Normalisasi Data pada Dataset	30
Gambar 4.15	<i>Splitting</i> Data <i>Training</i> dan Data <i>Testing</i>	30
Gambar 4.16	Pemodelan Algoritma <i>Linear Regression</i>	32



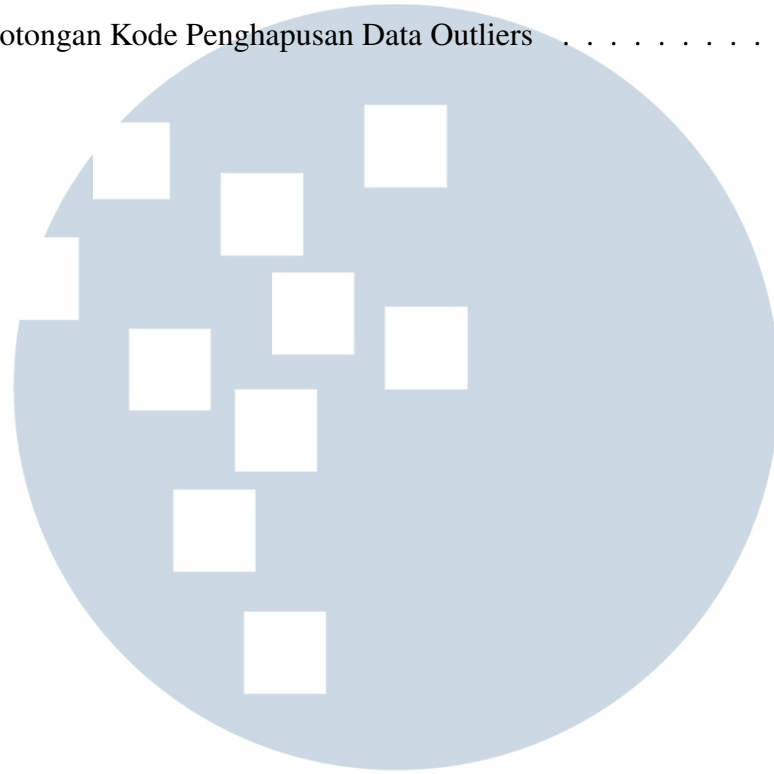
DAFTAR TABEL

Tabel 2.1	Deskripsi Variabel	8
Tabel 4.1	Deskripsi Variabel	15
Tabel 4.2	Tabel Heatmap Korelasi Antar Variabel	24
Tabel 4.3	Tabel Metriks Evaluasi	33



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

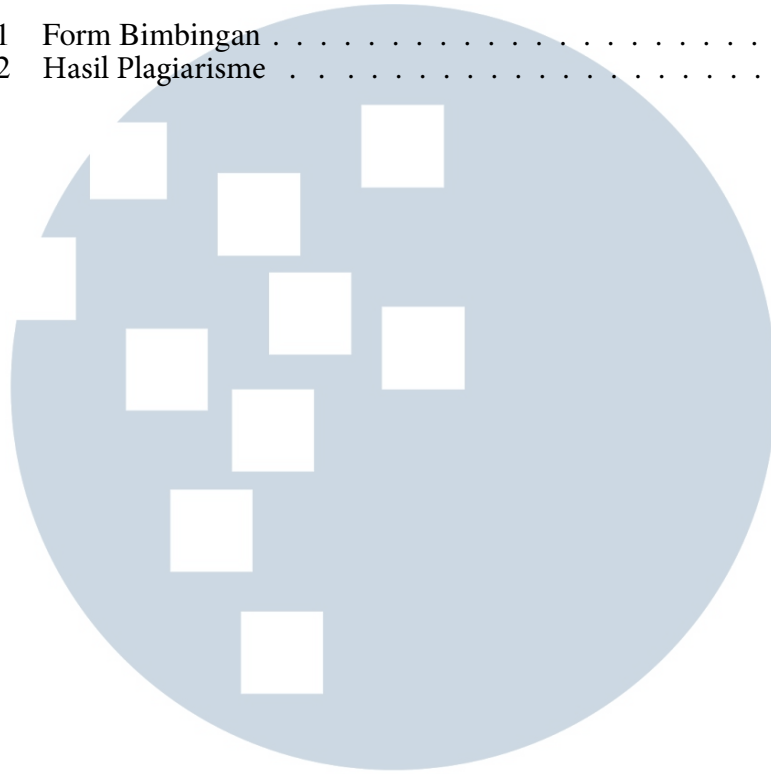
4.1 Potongan Kode Penghapusan Data Outliers 25



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran 1	Form Bimbingan	39
Lampiran 2	Hasil Plagiarisme	40



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA